

We claim:

- 1 1. An edge node adapted to serve a multiplicity of data streams, said edge node  
2 comprising:
  - 3 a plurality of input ports adapted to receive said data streams, wherein each  
4 of said input ports includes an input-port controller and at least one of said  
5 input ports includes a bitrate-estimation device adapted to compute a bitrate  
6 requirement for each of said data streams;
  - 7 a plurality of output ports, wherein each of said output ports includes an  
8 output-port controller;
  - 9 a switching fabric adapted to connect any of said input ports to any of said  
10 output ports; and
  - 11 an edge controller including:
    - 12 an edge control processor adapted to communicate with said plurality  
13 of input ports and said plurality of output ports;
    - 14 a route selection device, in communication with said edge control  
15 processor, adapted to select a route for each of said data streams;
    - 16 a fabric scheduling device, in communication with said edge control  
17 processor, adapted to determine a distinct time of transfer for each of a  
18 set of data segments, into which each of said data streams are  
19 segmented, across said switching fabric; and
    - 20 a bitrate-allocation device, in communication with said edge control  
21 processor, adapted to allocate a bitrate of a data stream based, at least  
22 in part, on said bitrate requirement for said data stream computed by  
23 said bitrate-estimation device.
- 1 2. The edge node of claim 1 wherein said switching fabric has a greater number of  
2 output ports than input ports.

- 1 3. The edge node of claim 1 wherein at least one of said output ports includes a  
2 time locking device adapted to control a data transmit time from said at least one of  
3 said output ports.
- 1 4. The edge node of claim 1 wherein said switching fabric is a common-memory  
2 fabric.
- 1 5. The edge node of claim 1 wherein said switching fabric is a space-switching  
2 fabric.
- 1 6. The edge node of claim 1 wherein said switching fabric is a rotator-based fabric.
- 1 7. The edge node of claim 1 wherein said switching fabric switches entire channels.
- 1 8. The edge node of claim 1 wherein said switching fabric switches data segments  
2 of time-shared channels.
- 1 9. The edge node of claim 1 wherein at least one of said plurality of input ports is a  
2 source port adapted to receive data from data sources and at least one of said  
3 plurality of input ports is a receiving port adapted to receive data from core nodes  
4 and other edge nodes.
- 1 10. The edge node of claim 1 wherein at least one of said plurality of output ports is a  
2 sink port adapted to transmit data to data sinks and at least one of said plurality of  
3 output ports is a departure port adapted to transmit data to core nodes and other  
4 edge nodes.
- 1 11. The edge node of claim 10 wherein each of said plurality of output ports is  
2 adapted to collate said data segments.
- 1 12. The edge node of claim 11 wherein said departure port is a time-locking  
2 departure port, said time-locking departure port including data buffers and time-  
3 locking circuitry.

1 13. The edge node of claim 1 wherein said bitrate-estimation device is adapted to  
2 base said bitrate requirement for each of said data streams on parameters  
3 associated with said each of said data streams.

1 14. The edge node of claim 1 where said bitrate-estimation device further comprises  
2 a hysteresis-control device adapted to base said bitrate requirement for each of said  
3 data streams on usage measurements associated with said each of said data  
4 streams.

1 15. The edge node of claim 14 wherein said input-port controller further comprises an  
2 input buffer and said hysteresis-control device comprises:

3 an on-off hysteresis-control unit adapted to:

4 determine an admission state of either accept or reject based on an  
5 occupancy of said input buffer; and

6 reject a request for a new connection for a data stream when said  
7 admission state is reject;

8 an incremental hysteresis-control unit adapted to request bitrate allocation  
9 increments.

1 16. A method of determining a bitrate requirement for a connection, said method  
2 comprising:

3 receiving a packet, where said packet includes an indication of a type for said  
4 connection;

5 determining, from said indication, a type for said connection;

6 if said connection is determined to be of a first type, extracting an explicit  
7 bitrate-allocation requirement from said packet;

8 if said connection is determined to be of a second type, computing said bitrate  
9 requirement from parameters characterizing said traffic associated with said  
10 connection; and

11 if said connection type is determined to be a of third type, deducing said  
12 bitrate requirement from data stream monitoring.

1 17. The method of claim 16 wherein said data stream monitoring comprises:

2 observing a buffer to obtain successive buffer-occupancy readings;

3 retaining at least two said buffer-occupancy reading;

4 comparing said at least two consecutive readings; and

5 determining a requirement for a bitrate allocation change based on said  
6 comparing.

1 18. The method of claim 17 further comprising modifying said deduced bitrate  
2 requirement based on said determining.

1 19. An edge node for high-speed traffic processing comprising:

2 a switching fabric;

3 a plurality of input ports in communication with said switching fabric, each of  
4 said plurality of input ports having an input-port controller;

5 a plurality of output ports in communication with said switching fabric, each of  
6 said plurality of output ports having an output-port controller; and

7 an edge node controller in communication with said switching fabric, each of  
8 said input-port controllers and each of said output-port controllers, said edge  
9 node controller including a high-speed scheduling device, said high-speed  
10 scheduling device operable to:

11 receive bitrate allocation information from at least one input-port  
12 controller;

13 maintain a state of a given input port associated with said at least one  
14 input-port controller;

15 maintain a state of each of said plurality of output ports;

16 assign time slots of a scheduling frame to communicate data segments  
17 from said given input port to one or more of said output ports, where  
18 the number of said time slots allocated to a given output port is based  
19 on said bitrate allocation information; and  
20 transmit said scheduling frame to said corresponding input-port  
21 controller.

1 20. The edge node of claim 19 wherein at least one of said output port controllers  
2 further includes a time-locking device for time-locking a corresponding one of said  
3 plurality of output ports to a core node.

1 21. The edge node of claim 19 wherein at least one of said input port controllers  
2 comprises a destination identifier adapted to translate an Internet Protocol address.

1 22. The edge node of claim 19 wherein at least one of said input port controllers  
2 includes a data formatter adapted to segment a received data stream into equal  
3 sized data segments.

1 23. The edge node of claim 19 wherein at least one of said input port controllers  
2 includes a bitrate allocation requirement estimator adapted to control requests for  
3 changes in bitrate allocation.

1 24. The edge node of claim 19 wherein said input-port controller is adapted to  
2 receive a data stream that specifies a destination and an explicit bitrate allocation  
3 requirement and wherein said edge node controller is adapted to provide a route to  
4 said destination having said explicit bitrate allocation requirement.

1 25. The edge node of claim 19 wherein said input-port controller is adapted to  
2 receive a data stream that specifies a destination and estimate a bitrate allocation  
3 requirement for said data stream and wherein said edge node controller is adapted  
4 to provide a route to said destination having said estimated bitrate allocation  
5 requirement.

1 26. The edge node of claim 25 wherein said input-port controller is adapted to

2 estimate a requirement for bitrate allocation according to a hysteresis-control  
3 mechanism; and

4 transmit a request for a change in said estimated bitrate allocation  
5 requirement for said data stream.

1 27.A method of data scheduling in an edge node comprising:

2 receiving bitrate allocation information from an input-port controller;

3 maintaining a state of a given input port associated with said input-port  
4 controller;

5 maintaining a state of each of a plurality of output ports;

6 assigning time slots of a scheduling frame to communicate data segments  
7 from said given input port to one or more of said output ports, where the  
8 number of said time slots allocated to a given output port is based on said  
9 bitrate allocation information; and

10 transmitting said scheduling frame to said input-port controller.

1 28.A device for high-speed scheduling operable to:

2 receive a bitrate allocation from an input-port controller;

3 generate a bitrate demand matrix from said bitrate allocation;

4 where a schedule is defined by a plurality of time slots, assign said time slots  
5 to input-output pairs according to said bitrate demand matrix, an input state  
6 matrix and an output port state matrix; and

7 transmit said schedule to said input-port controller.

1 29.A method of scheduling a transfer of data segments through a switching fabric  
2 from a plurality of input ports of said switching fabric to a plurality of output ports of  
3 said switching fabric wherein a scheduling frame having a predetermined number of

4 time slots and an equal, predetermined duration is associated with each of said  
5 plurality of input ports, said method comprising:

6 dividing said duration of each said scheduling frame into a plurality of time  
7 windows;

8 dividing said plurality of input ports into input-port groups;

9 for each said scheduling frame, forming a plurality of non-intersecting  
10 scheduling domains, where each said scheduling domain associates one of  
11 said input-port groups with one of said plurality of time windows;

12 concurrently executing scheduling processes, one scheduling process for  
13 each of said non-intersecting scheduling domains, each said scheduling  
14 process determining a schedule that indicates, for a given scheduling domain,  
15 a timing for a transfer of data segments to selected ones of said plurality of  
16 output ports from said input ports associated with said time window in said  
17 given scheduling domain; and

18 at least where said executing does not use all said time windows of a given  
19 scheduling frame, repeating said forming and said executing for permutations  
20 of said input-port groups and time windows.

1 30. The method of claim 29 wherein, for a scheduling frame associated with a given  
2 input port, said executing scheduling process includes allocating at least one time  
3 slot to an output port to which said given input port has non-zero traffic.

1 31. The method of claim 30 wherein said executing said scheduling process for a  
2 given scheduling domain comprises sequentially executing said scheduling process  
3 for said input ports of said input-port group of said given scheduling domain.

1 32. The method of claim 29 further comprising receiving, from each of said plurality of  
2 input ports, a demand vector, where said demand vector indicates a required  
3 number of time slots per scheduling frame for each of said plurality of output ports  
4 and wherein said demand vector is structured as an aggregate-demand ring-list  
5 including a plurality of data records, each said data record including a field indicating

6 a preceding record identifier, a field indicating a succeeding record identifier, a field  
7 indicating an output port identifier, and a field indicating a non-zero number of  
8 required time slots per scheduling frame.

1 33. The method of claim 29 further comprising receiving, from each of said plurality of  
2 input ports, a demand vector, where said demand vector indicates a required  
3 number of time slots per scheduling frame for each of said plurality of output ports  
4 and wherein said demand vector is structured as a flat-demand ring-list including a  
5 plurality of data records, each said data record including a field indicating a  
6 preceding record identifier, a field indicating a succeeding record identifier, and a  
7 field indicating an output port identifier.

1 34. The method of claim 29 further comprising limiting an execution time period for  
2 each said scheduling process.

1 35. The method of claim 34 wherein said limited execution time period varies with  
2 each of said repeating of said concurrently executing a scheduling process.

1 36. The method of claim 29 wherein said switching fabric is a space switch.

1 37. The method of claim 29 wherein said switching fabric is a rotator-based switch  
2 having a plurality of middle memories.

1 38. The method of claim 37 wherein, for each said input port, a time slot is uniquely  
2 associated with one of said plurality of middle memories.

1 39. The method of claim 29 further comprising using said schedule, determined for  
2 said duration of said scheduling frame, for a configuration period, where the duration  
3 of said configuration period is an integer multiple of said duration of scheduling  
4 frame.

1 40. In an edge node comprising a plurality of input ports, a plurality of output ports  
2 and a switching fabric, a fabric-scheduling apparatus comprising:

3 a plurality of input-port scheduling devices, where each of said input-port  
4 scheduling devices is associated with an input-port group and each input-port



group is a sub-set of said plurality of input ports, each of said input-port scheduling devices including:

a receiver adapted to receive bitrate allocations from controllers of said plurality of input ports and translate each said bitrate allocation into translated bitrate allocations, where each said translated bitrate allocation indicates a required number of time-slots for a predefined scheduling frame duration;

a ring-list generator adapted to form a ring-list from said translated bitrate allocations received from said receiver;

a ring-list time-slot assignment circuit adapted to:

receive said ring-list from said ring-list generator;

perform an assignment of each time-slot in a predefined time-slotted scheduling frame to a transfer of a data segment from one of said input ports in said input-port group associated with said input-port scheduling devices to one of said plurality of output ports; and

modify said ring-list to reflect said assignment;

a ring-list memory adapted to maintain said modified ring-list; and

a plurality of output-state memory devices, each of said output-state memory devices associated with a predefined time window; and

an access-control device adapted to cyclically connect each of said input-port scheduling devices with each of said output-state memory devices.

41. The edge node of claim 40 wherein said access control device comprises a two-way rotating connector.

42. The edge node of claim 41 wherein said ring-list generator generates an aggregate-demand ring-list.

- 1 43. The edge node of claim 41 wherein said ring-list generator generates a flat-
- 2 demand ring-list.

14675ROUS01U 56